View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Language Science Press

Chapter 1

Linguistic complexity: A case study from Swahili

Kyle Jerro
University of Essex

This paper addresses the question of linguistic complexity in Swahili, a Bantu language spoken in East and Central Africa. Literature on linguistic complexity in other languages has argued that high levels of second-language learning affect linguistic complexity over time. Swahili serves as an ideal case study for this question because it has been used as a lingua franca for several centuries. I compare the phonological and morphological systems in Swahili to five other related Bantu languages, as well as compare all six languages to the original Proto-Bantu systems. The results of the study show that there is no decrease in phonological or morphological complexity in (standard) Swahili when compared to other closely related Bantu languages, though the grammar has strongly diverged from the other related languages.

1 Introduction: the question of linguistic complexity

It is generally assumed by linguists that all languages share the same level of complexity, with "simpler" areas of grammar being compensated by more complexity elsewhere. Some researchers take this as a core design feature of language (cf. work from the generative perspective, such as Pinker & Bloom 1990; Pinker 1994; Baker 2003), though this has tacitly pervaded most linguistic thought.

Recently, however, work by various linguistic typologists has put this assumption into question, investigating several linguistic domains (see Miestamo 2008; Sampson 2009 and Givón & Shibatani 2009 for overviews of the literature on complexity). A core area of the research in this field is simply how to answer such a question (Nichols 2009; Sampson 2009; Miestamo et al. 2008). For example, Nichols (2009) compares various features of languages, such size of phoneme inventory, number of inflectional categories on a basic verb, number of alignments in a single language, etc. Other work situates linguistic complexity within a social context. One claim is that older languages tend to be more complex that new ones (e.g. Creoles), cf. McWhorter (2008) and Trudgill (2009).

Kyle Jerro. Linguistic complexity: A case study from Swahili. In Jason Kandybowicz, Travis Major, Harold Torrence & Philip T. Duncan (eds.), *African linguistics on the prairie: Selected papers from the 45th Annual Conference on African Linguistics*, 3–19. Berlin: Language Science Press. DOI:10.5281/zenodo.1251708

Another claim is that population size relates to linguistic complexity (Trudgill 2004; Hay & Bauer 2007; Nichols 2009).

Another vein of this literature – and the topic of this paper – has investigated the interaction of complexity and language contact, claiming that high amounts of second-language learning, including the use as a lingua franca, affects linguistic complexity and increases the rate of language change (Kusters 2003a,b; Trudgill 2009; McWhorter 2008; 2011; Trudgill 2011). Trudgill (2011) claims that that the specific effect on complexity is contingent upon the nature of second-language learning: while large amounts of second-language learning by adult speakers may result in net decomplexification, learning by children (e.g. through prolonged contact between two languages) may lead to *increased* complexity. This paper tests the affects of language contact on complexity in Swahili, used as a lingua franca throughout much of East and Central Africa. I compare Standard Swahili to neighboring Bantu languages in their synchronic morphological and phonological features as well as their divergence from Proto-Bantu.

To test this claim, I employ similar metrics of complexity to those used by Kusters and McWhorter (i.e morphology, see §5), comparing different aspects of Swahili morphology to the grammar of five sister languages. In addition, I discuss the phonological inventories of the languages, a component absent from Kusters' and McWhorter's studies, but discussed at length by others (Hay & Bauer 2007; Trudgill 2011). From the comparisons, I conclude that Swahili does not exhibit any systematic decomplexification in comparison to the other languages, though it shows several grammatical differences from related languages. This situation is predicted from the framework proposed in Trudgill (2011), where long-term bilingualism (here, between Swahili and Arabic) may lead to the rapid change of a contact language.

The remainder of this paper is organized as follows: in §2, I summarize the claims of the decomplexification hypothesis. I then outline the linguistic and sociolinguistic situations of five Bantu languages from East Africa chosen to serve as comparison cases. Sections 4-5 use phonological and morphological metrics, respectively, in order to compare the complexity of Swahili to the comparison languages. Section 6 discusses the findings and their relation to the the decomplexification hypothesis.

2 Contact and (de-)complexification

In research on complexity, two opposite effects on complexity have been found, depending on the nature of the linguistic community. Languages in prolonged contact regions tend to develop high amounts of linguistic complexity (Heine & Kuteva 2005; Dahl 2004; Givón 1984). On the other hand, situations with high numbers of sudden second-language learners result in simplification of linguistic structure. As discussed in Trudgill (2011), the crucial divide between the two groups is the critical period of language acquisition: adult learners are not as adept as children at acquiring a (second) language. In a situation where adult speakers are acquiring a language, this "sub-optimal acquisition" (a term from Dahl 2004) results in the reduction of ornamental or non-obligatory elements of grammar.

As Kusters (2003b) states, "the more second-language learning has taken place in a speech community, the more internal dialect contact and migrations occurred, and the less prestige a language has, the more *transparent* and *economic* the verbal inflection will become" (275, emphasis in original). For Kusters, an inflectional system is more economic if it makes fewer category distinctions. In order to test the prediction of the decomplexification hypothesis, lingua francas that have been used by many second-language learners can be compared to sister languages or varieties that have not been used as lingua francas.

Kusters (2003b,a) provides several case studies in contact languages that have undergone decomplexification, tracing the changes from an older stage of the language to various modern sister languages. For example, one case study comes from three descendants of Old Norse: Icelandic, Faroese, and Standard Norwegian. He argues that the varieties that are more insular have maintained complexity that is absent in metropolitan varieties (i.e. the dialect of the capital city of the Faroese Islands, Tórshavn). As an example, consider the data in Table 1, with the verb forms for the verb 'to awake' in Old Norse and three descendant languages (Kusters 2003b: 285, Table 5).

Table 1: Verbal tense in Old Norse and descendant languages

	Old Norse	Icelandic	Faroese	Tórshavn
1sg	vakn-a	vakn-a	vakn-i	(-′)
2sg	vakn-ar	vakn-ar	vakn-ar	(-′r)
3sg	vakn-ar	vakn-ar	vakn-ar	(-′r)
1pl	vakn-um	vökn-um	vakn-a	(-′)
2pl	vakn- i_	vakn-ið	vakn-a	(-′)
3pl	vakn-a	vakn-a	vakn-a	(-′)

He argues that Faroese, a variant that has been in prolonged contact with Danish, has reduced morphological complexity from the Old Norse, and Tórshavn has undergone further reduction, having only stress as a indicator of tense. The only person marking is the marking of second- and third-singular, to the exclusion of all other persons and numbers. In addition, the Tórshavn dialect has completely neutralized certain inflectional categories, like past indicative and present subjunctive.

McWhorter (2011; 2008) makes the stronger claim that second-language learning is the *only* factor that drives overall simplification in a language. Namely, sweeping loss of complexity in a language is impossible without the influence of second-language learning. The argument works in the opposite direction from Kusters'; when you find an instance of decomplexification, it is predicted that this must have come from a situation of high second-language learning. McWhorter's metrics of complexity are similar to those of Kusters (2003b). For example, in his 2008 paper, he compares two varieties of the Tetun language spoken in Timor. The first, Tetun Dili, is used as a lingua franca by

two-thirds of the island; the other, Tetun Terik, is only spoken on the southern coastline. McWhorter predicts that because Tetun Dili is a lingua franca, it has a simpler grammar than Tetun Terik. He presents several instances where the Dili variety is more economical in the number of morphological categories it has. For example, while Terik has three verbal affixes, Dili has two; Tetun has six numeral classifiers while Dili only has four (and those four are used optionally); Tetun has an overt marker for definiteness, while Dili uses context to indicate this; Tetun has three copulas, while Dili has only one; etc. In short, the variety that is used as a lingua franca is systematically simpler than a sister variety without the same level of second-language use.

When two languages are in prolonged contact, and the acquirers of a second language are mostly children, the opposite effect is found: over time, more complexity is found, often by the additive borrowing from the neighboring language. For example, Comrie (2008) and Trudgill (2011) cite the example of Michif, a mixed language from contact between Cree and French (Bakker 1997). Michif, from prolonged multilingualism with French and Cree, developed an elaborate grammar, taking grammatical elements from both Cree and French, with verbal structure inherited from the former and nominal structure from the latter. The result is that Michif employs elaborate verbal and morphological categories found in neither French nor Cree.

In short, work on contact and complexity has found three related effects of contact: first, language contact increases the rate of language change; second, second-language learning by adults often leads to reduction in complexity via imperfect acquisition; and, third, prolonged contact between two languages often results in complexification as forms are taken from one and added into the other. In this paper, I tease apart the level of complexity of standard Swahili, comparing it to five related Bantu languages that have not had parallel situations of language contact.

3 Swahili and the five comparison languages

Swahili serves as another ideal case study in fleshing out the claims of the decomplexification hypothesis. Swahili is spoken as a native language along the Indian Ocean coast of Kenya and Tanzania and in the Zanzibari archipelago. It is also used as an official language and lingua franca in Kenya, Tanzania and the Democratic Republic of the Congo (DRC) in addition to a language of business and commerce at different points in history in Uganda, Rwanda, and Burundi. Because of this widespread use as a lingua franca, nearly 140 million people use Swahili as a second language, while only 5 million speak it natively. Given the overwhelming predominance of second-language speakers of the language, the decomplexification hypothesis predicts that Swahili should be systematically less complex than related languages with little or no use by second-language speakers.

I have chosen five languages spoken in the countries where Swahili is or has been routinely used as a lingua franca. I have chosen one language from each country, and the languages are all part of the Northeastern branch of the Bantu family (with the exception

of Lingala).[1] The comparison languages are Gikuyu (Kenya, E.51), Lingala (DRC, C.30B), Haya (Tanzania, JE.22), Kinyarwanda (Rwanda, DJ.61), and Luganda (Uganda, JE.15).

Gikuyu is spoken in Central Kenya by the Gikuyu people, numbering at approximately 7 million. Lingala is a language spoken by approximately 2 million people in the Republic of Congo, the Democratic Republic of Congo, and parts of the Central African Republic. Haya is spoken in Northwestern Tanzania, near the shores of Lake Victoria (Byarushengo et al. 1977). There are approximately 1 million speakers of the language. Luganda is spoken by approximately 4 million people in Southern Uganda. Though used mostly by the Baganda people, it is also used as a second language by approximately 1 million people in Uganda (Ethnologue 2013). Although the use of Luganda by second-language learners is not ideal as a comparison case in the current study, the situation of Luganda is different from Swahili in that the majority of speakers use Luganda as a first language. Swahili on the other hand, is used overwhelmingly as a second language. Kinyarwanda is spoken by somewhere around 12 million people in Rwanda, Burundi, and parts of Uganda and DRC.

4 Phonological complexity

The first metric I use to compare the relative complexity among these languages is their phonological inventories. Phonological complexity did not figure in Kusters' and McWhorter's discussions, though several other works have used phonological inventory as a metric for calculating complexity (Hay & Bauer 2007; Nichols 2009). The decomplexification hypothesis as outlined above predicts that Swahili will have the smallest inventory of phonemes; over time, imperfect learning by second-language speakers would result in the reduction of phoneme contrasts not found in their first languages. Over time, this reduced vowel inventory becomes the standard inventory of the language.

4.1 Vowel complexity

4.1.1 Vowel inventory

Bantu languages generally have between five and seven vowels in their inventory, and they generally include tonal and length distinctions (Hyman 2003; Maddieson 2003). Proto-Bantu has been reconstructed to have seven vowels with high and low tone contrasts. Table 2 indicates the number of different vowels (based on quality) in each of the languages in the test set as well as whether each language makes a distinction between long and short vowels and between tones.

Numerically, Swahili has a simpler vowel inventory than the other languages; it has two fewer vowels than Proto-Bantu. Furthermore, Swahili has lost the tone and length

[1] A better comparison set may be languages that are more closely related to Swahili genetically than the five chosen here. Accessibility to resources was a major factor in linguistic choice, though the localization of these languages to East Africa is intentionally aimed at keeping to languages that are more similar to Swahili.

Table 2: Size of vowel inventories

Language	Vowels	Tone	Length	Source
Proto-Bantu	7	+	+	Maddieson (2003)
Swahili	5	–	–	Ashton (1966)
Gikyuyu	7	+	+	Barlow (1960)
Lingala	7	+	+	Guthrie (1966)
Haya	5	+	+	Byarushengo et al. (1977)
Kinyarwanda	5	+	+	Myers & Crowhurst (2006)
Luganda	5	+	+	Kirwan & Gore (1951)

contrasts in Proto-Bantu, while the other languages have retained these features. This is the kind of inventory reduction expected by the decomplexification hypothesis.

4.1.2 Other kinds of vowel complexity

Although the size of vowel inventories indicates a lower level of complexity in Swahili, another possible metric is linguistic markedness (cf. McWhorter 2008; 2011). Swahili, unlike its sister languages, shows three linguistically marked phonological processes that are absent in the other languages. These processes include the permission of syllabic consonants, an irregular stress system, and vowel hiatus. Unlike a numerical metric like phoneme inventory, however, phonological operations in a language are not as easily quantifiable. However, I argue here that the quantitatively fewer phonemic vowel contrasts in Swahili are counteracted by the complexity that ensues with respect to its vowel system.

First, Swahili has syllabic nasal consonants (Ashton 1966). This is present on words such as *mtoto* [m.toto] 'child,' *mtu* [m.tu] 'person,' and *mlango* [m.lango] 'door.' Of the sister languages, only Haya permits syllabic consonants; all maintain a minimal (C)CV syllable structure (cf. the cited grammars). Interestingly, Hyman (2003) assumes this is a natural change, derived from the loss of [u] in *mu-* nominal prefixes.

A further noteworthy difference between Swahili to the exclusion of the other languages is that Swahili permits vowel hiatus, with juxtaposed vowels serving as nuclei of separate syllables. For example, *chui* 'leopard' is syllabified as [tʃu.i], and *paa* 'gazelle' as [pa.a]. The other languages do not permit vowel hiatus; Kinyarwanda, for example, deletes one of any two adjacent vowels, even between word boundaries. For example, the sentence *uri umwana* 'you are a child' is pronounced [u.ru.mŋa.na], with the word-final [i] in *uri* being deleted.

Finally, unlike the other languages of the study, Swahili has several cases of irregular lexical stress.[2] In most Bantu languages, stress falls on the penultimate syllable. In Swahili, however, there are cases where Arabic loanwords carry stress on the antepenul-

[2]Thanks to Scott Myers for suggesting this point.

timate syllable, in words such as *nusura* ['nu.su.ra] 'almost,' *ratili* ['ra.ti.li] 'pound,' and *thumuni* ['tʰu.mu.ni] 'an eighth' (Ashton 1966). Here, contact with Arabic is the obvious influence of the complexification of the Swahili stress system.

These three examples show that despite the smaller phonemic inventory, Swahili has elements of complexity that are absent in the other languages. These features, however, are difficult to quantify, and their inclusion in metrics of complexity vary. My conclusion from the data in this section is that there is no clear reduction in complexity in the vowel system of Swahili.

4.2 Consonant inventory

Although the number of vowels in Swahili is quite low, the consonant inventory is noticeably larger than the inventories of the comparison languages.[3]

Table 3: Size of consonant inventories

Language	Consonants
Proto-Bantu	11
Swahili	30
Gikuyu	14
Lingala	15
Haya	19
Kinyarwanda	22
Luganda	18

The consonant inventory in Swahili is striking larger than the other languages under discussion, being over two times larger than the consonant inventory of Gikuyu and Proto-Bantu.[4] The larger inventory in Swahili comes in part from having both voiced and voiceless stops and fricatives for bilabial, alveolar, and velar places of articulation. Many languages lack a subset of these sounds, often having only the voiced or voiceless counterpart. Gikuyu, for example, lacks the voiceless bilabial stop, the voiceless velar fricative, and the voiced alveolar fricative that are found in Swahili.

A further difference is that Swahili is the only language in the group with the aspirated stops and fricatives [pʰ tʰ tʃʰ kʰ] (Ashton 1966; Engstrand & Lodhi 1985). Aspiration is also found in various other Bantu languages, such as Zulu, Swati, Makua, Doko, Chichewa, and Kongo. It has been argued that aspiration is a possible outgrowth

[3] The inventories in Table 3 come from the same sources as in Table 2, save for the number for Proto-Bantu, which comes from Hyman (2003).

[4] Nasalized consonants were not counted for any of the languages, as the descriptions of them were not satisfactorily convincing that these were indeed separate phonemes. The inclusion of these sounds in the data would not affect the trend, however, since they are also a class of sounds reported in Swahili.

of a consonant followed by the Proto-Bantu high vowels (Hyman 2003) or from an earlier voiceless pre nasalized stop (Maddieson 2003). Regardless of the origin of phonemic aspiration, the presence of aspiration results in a notable increase in the phonemic inventory of Swahili, resulting in a larger inventory than the comparison languages, as well as an innovation since Proto-Bantu.

Another interesting feature of the Swahili consonant system is that all voiced stops are implosives. Swahili has four of these phonemes: [ɓ ɗ ʄ ɠ]. Implosive stops are not found in any of the comparison languages from East Africa, though implosive stops are documented in the southern Bantu languages, with Maddieson (2003) treating implosives in the Bantu family as a natural development in some daughter languages.

4.3 Discussion

The decomplexification hypothesis predicts that Swahili should have a noticeably smaller phoneme inventory than the comparison languages. Although this is true with vowel inventory, the consonant inventory in Swahili is markedly larger than any of the other comparison languages. Importantly, the Swahili consonant system is nearly three times larger than in Proto-Bantu, suggesting considerable innovation during the evolution of Swahili.

5 Morphological complexity

The next domain of investigation is the morphological (dis)similarity between Swahili and the other Bantu languages. If the decomplexification hypothesis is correct, it is expected that Swahili will make fewer distinctions and that morphemes will be more phonologically reduced than the other languages. I investigate the domains of noun class morphology, valency-changing morphology, and tense/aspect/mood morphology, which are all three morphological domains that are found in each of the languages.

5.1 Gender classes on nominals

Bantu languages are well known for their rich noun class morphology. The noun classes for Swahili, Haya, Kinyarwanda, Luganda, and Lingala are provided in Table 4, as well as the reconstructions of the Proto-Bantu inventory (Meeussen 1967; Schadeberg 2003a).[5] Given then decomplexification hypothesis, it is expected that Swahili should be more economic in its morphological forms, either in the phonological shape of the morphemes or in the number of semantic distinctions.

Swahili has a comparable number of category distinctions to the other languages; although it is reduced from Proto-Bantu, only one of the other languages retains the number of category distinctions found in Proto-Bantu (i.e. Luganda). Clearly, the prediction

[5]The source for Gikuyu did not include enough detail for this comparison. The sources for the modern languages in Table 4 are: Swahili (Ashton 1966), Haya (Byarushengo et al. 1977), Kinyarwanda (kinyarwanda.net), Luganda (Kirwan & Gore 1951), and Lingala (Guthrie & Carrington 1988).

Table 4: Comparison of noun class morphology

Class	Swahili	Haya	Kinyarwanda	Luganda	Lingala	PB
1	m(u)-	mu-	umu-	(o)mu-	mo-	*mu-
2	wa-	ba-	aba-	(a)ba-	ba-	*ba-
3	m(u)-	mu-	umu-	(o)mu-	mo-	*mu-
4	mi-	mi-	imi-	(e)mi-	mi-	*mi-
5	ji-	li-	iri-	li-, eri-	li-	*ị-
6	ma-	ma-	ama-	(a)ma-	ma-	*ma-
7	ki-	ki-	iki-	(e)ki-	e-	*ki-
8	vi-	bi-	ibi-	(e)bi-	bi-	*bị-
9	n-	n-	i(n)-	(e)n-	N-	*n-
10	n-	n-	i(n)-	(e)n-	N-	*n-
11	u-	lu-	uru-	(o)lu-	lo-	*du-
12	n-	ka-	aka-	(a)ka-	bo-	*ka-
13	-	tu-	utu-	(o)tu-	-	*tu-
14	-	bu-	ubu-	(o)bu-	bo-	*bu-
15	ku-	ku-	uku-	(o)ku-	ko-	*ku-
16	pa-	-	aha-	wa-	-	*pa-
17	ku-	-	-	ku-	-	*ku-
18	mu-	-	-	mu-	-	*mu-
19	-	-	-	-	-	*pị-
20	-	-	-	(o)gu-	-	-
21	-	-	-	-	-	-
22	-	-	-	(a)ga-	-	-
23	-	-	-	e-	-	*i-
	16	15	16	21	14	21

that Swahili exhibit a noteworthy reduction in the number category distinctions is not borne out in this comparison.

As for the phonological shape of the morphemes, Swahili lacks the pre-prefix that is found in Luganda and Rwanda. At a first glance, this could be argued to be an instance of phonological reduction in Swahili. However, it has been argued in the literature that these pre-prefixes were not present in Proto-Bantu (Katamba 2003), suggesting that the pre-prefix in languages that have it is an innovation.

Support for this point is that the use of the pre-prefix varies drastically in the languages which use it. In Luganda, a variety of features converge to predict the presence of the pre-prefix, such as whether the noun is a dependent or main clause, appears in the affirmative or negative, etc. (Hyman & Katamba 1991; 1993). In Zulu, it has been argued that the pre prefix is a case marker for nominals that lack structural case (Halpert 2012). Zerbian & Krifka (2008) show that features such as genericity, specificity, and def-

initeness are present in various languages which utilize the pre-prefix, such as Xhosa, Bemba, and Kinande. Crucially, it is assumed that the pre-prefix is a later innovation from Proto-Bantu, perhaps being a reanalysis of cliticized pronouns onto the main noun (Bleek 1869).

The lack of a pre-prefix in the Proto-Bantu stems, as well as the semantic nature of pre-prefixes in the languages which have them, suggests that the reduced phonological shape of class morphology in Swahili is not driven by phonological reduction due to second-language learning. Instead, Swahili has retained the original shape of Proto-Bantu stems.

5.2 Valency-changing morphology

Bantu languages utilize morphology to indicate valency changes to the argument structure of a verb. Both argument-adding (applicatives and causatives) and argument-reducing (stative, reciprocal, passive) morphology is employed by these languages. If the decomplexification hypothesis is correct, it is expected that valency-changing morphology in Swahili is simpler than in the comparison languages – be it phonologically reduced or with fewer morphological category distinctions.

Table 5 gives the morphological forms for different valency-changing morphology in Swahili (Russell 2003), Lingala (Guthrie 1966), Kinyarwanda[6] (Jerro 2015), Haya (Byarushengo et al. 1977), and the reconstructed forms in Proto-Bantu (Schadeberg 2003b).[7]

Table 5: Comparison of valency-changing morphology

Type	Swahili	Lingala	Kinyarwanda	Haya	PB
Benefactive	-(l)e /-(l)i	-el	-ir/-er	-il/-el	*-ɪl
Instrumental	-(l)e /-(l)i	-	-ish/-esh	-is/-es	*–ɪl
Locative	-(l)e /-(l)i	-	-ir/-er	-il/-el	*-ɪl
Causative	-ish/-esh	-is	-ish/-esh	-is/-es	*-i/-ici
Stative	-ik/-ek	-an	-ik/-ek	-ek	*-ɪk
Reciprocal	-an	-an	-an	-aŋgan	*-an
Passive	-(li)w/-(le)w	-	-w	-w	*-ʊ/-ɪbʊ

The first three types of morphology are applicatives, which add a new object to the valency of a verb. Reciprocals, statives, and passives all decrease the valency of a verb by one: reciprocals link the action back to the subject, i.e. the subject does the action to him or herself; passives demote the subject to an oblique position and promote the object to subject position; and statives describe the result state of a transitive verb.

[6] Those familiar with Kimenyi (1980) will notice that the locative applicative morpheme for Kinyarwanda in Table 5 differs from Kimenyi's description. Jerro (2015) describes a different locative applicative form for his speakers, who find Kimenyi's locative applicatives ungrammatical.

[7] The resources for Gikuyu and Luganda do not explicitly discuss valency-changing morphology.

Contrary to the decomplexification hypothesis, the data in Table 5 show that Swahili does not have a simpler system of valency-changing morphology. From the perspective of the number of category distinctions, it has a comparable number to the other languages, and has lost no form reconstructed for Proto-Bantu.

From the perspective of the phonological shape of the morphemes, there is no evidence that Swahili is simpler than the other languages. Many of the valency-changing forms in Bantu undergo vowel harmony with the nearest stem on the vowel, and Swahili is not an exception to this; it employs vowel harmony on valency-changing morphology in the same way as its sister languages.

In fact, if any argument were to be made regarding the complexity of valency-changing morphemes, Swahili is more complex in the phonological shape of its passive morpheme, which varies by context depending on the phonological shape of the verb to which it is applied (Russell 2003). The most productive form of the Swahili passive is *–w*, as in *fung–w–a* from *funga* 'fasten' and *tumi–w–a* from *tumia* 'use.' When the verb stem ends in [o] or [e], the form *–lew* is used. If the verb stem ends in [a] or [u], the form *–liw* is used, as in *za–liw–a* from *zaa* 'give birth' and *fu–liw–a* from *fua* 'wash clothes'. Russell (2003) also notes that the passive forms *–ew* and *–iw* are used with verbs of Arabic origin, such as *sameh–ew–a* from *samehe* 'forgive' and *hitaj–iw–a* from *hitaji* 'need.' In short, to form a passive in Swahili, there are complex factors that determine the phonological shape of the passive morpheme, and these factors are not present in the comparison languages.

In Kinyarwanda and Haya, for example, the passive form is *–w* for all verbs, and Lingala lacks a separate passive morpheme altogether (Guthrie 1966). This is evidence that valency-changing morphology in Swahili is not simpler than the sister languages, and in the domain of the passive, Swahili is actually more complex than the other forms.

5.3 Tense, aspect, and mood

Bantu languages have rich systems of tense, aspect, and mood (TAM). From the view of complexity, there are two ways in which a language may be simpler than the others with respect to TAM morphology. The language could make fewer distinctions in its tense, aspect, and mood categories, leaving TAM information to pragmatics. Another indication of decomplexification is if the language shows phonological reduction of the forms compared to other languages or from the protolanguage.

In Bantu languages, aspect and mood morphology generally appears as a prefix before the verb stem, but after the agreement subject marker. Aspect, on the other hand, appears as a suffix after the verb stem. If a language marks subjunctive or indicative, this appears in the aspect slot. The general template for TAM on a verb in Bantu is given in (1) (cf. Meeussen 1967; Nurse 2003).

(1) Subject Marker – Tense – STEM – Aspect/Subjunctive

Table 6 includes data for five different kinds of TAM that are prevalent in Bantu languages: tense, indicative/subjunctive, aspect, negation, and idiosyncratic TAM morphology that does not fit consistently with the other categories.[8]

[8]Data from Lingala and Haya are not included due to a paucity of description of the tense/aspect systems in those languages.

Table 6: Comparison of tense, aspect, and mood morphology

Type	Swahili	Luganda	Gikuyu	Rwanda
Present	na-	∅	∅	∅
Present II	a-	-	-	-
Pres. Continuous	-	-	ra-	ra-
Recent Past	li-	a-	∅	a-
Distant Past	-	MS	a-…-ire	ara-
Perfect	me-	-	-a	-
Past Perfect	-	-	-ite	-
Immediate Future	ta-	naa-	kũ-	za-
Near Future	-	li-	ka-	-
Distant Future	-	-	rĩ-	-
Imperative	-e	-e	-e	-e
Subjunctive	-e	-e	-e/-(n)i	-e
Indicative	-a	-	-	-
Imperfective	-	-	-ga	-a(ga)
Perfective	-	-	-a	-(y)e
Negation	hu-/si-	si-	ti-	si-
Conditional	nge-/ngali-	andi-	ngĩ-	ni-
Habitual	hu-	-	ga-	-
Narrative	ka-	ne-	-	-
‘not yet’	ji-	naa-	-	-
‘even if’	japo-	-	-	-
‘if’	ki-	-	-	-
‘still’	-	kya-	-	-
optative	-	-	ro-	-
‘also’	-	-	-	na-
	15	12	17	12

The first section shows various tense morphemes: variants of past, present, and future. For some languages (such as Gikuyu and Kinyarwanda) there are various past and future tenses, depending on the temporal proximity to the speech event. For languages with only one distinction for a particular tense, the form is listed in the tense closest to the present. For example Swahili only has one past tense, which is listed in the “Recent Past” row. The abbreviation MS for Luganda, indicates that a “modified stem” is used to indicate the distant past, formed by a lexically-determined set of stem-changing operations. In Gikuyu, the distant past is marked by the combination of a prefix and suffix, indicated by *a-…-ire*. In Swahili, the present *na-* can also be used for present continuous.

The second category covers indicative, subjunctive, and imperative morphology, found consistently among all of the languages. For Swahili, the final vowel *-a* is used as a general indicative mood marker.

The third category is aspect. Kinyarwanda and Gikuyu both have a distinction between perfective and imperfective, while Swahili and Luganda do not have morphology for these aspectual distinctions.

All of the languages share cognate morphology for negation.

Other mood distinctions are covered in the final section of Table 6. This is reserved for mood categories that are highly idiosyncratic meanings in particular languages, such as morphology for meanings such as 'not yet' and 'still' in Swahili and Luganda, respectively. Another is the "optative" in Gikuyu, used for blessings and curses (Barlow 1960). The narrative morpheme is used for verbs that are in a series during a narration of events.

There is no clear indication that any of these languages has a notably simpler system of TAM morphology. Summing the number of morphological category distinctions made in the four languages, it is clear that the inventory of distinctions is quite comparable for all the languages, and Swahili is not noticeably less complex than any other language. It is important to note the heterogeneity among the languages' TAM morphology; few morphemes are cognate, which makes it impossible to compare the phonological reduction among the languages, meaning that the phonological reduction of forms cannot be measured for complexity.

6 Discussion: complexity and language contact

Data comparing the phonological inventory and morphological systems among Swahili, Gikuyu, Kinyarwanda, Lingala, Haya, and, Luganda – as well as a comparison with Proto-Bantu – show that there is no instance where clear decomplexification has occurred in Swahili. In fact, in some instances, such as in consonant inventory, Swahili shows more complexity that the other languages. In nearly all of the grammatical properties discussed, Swahili is highly divergent from the other languages, with notable differences in phonological inventory, such as a larger consonant inventory, a smaller vowel inventory, and irregularities with respect to stress and syllabification. Crucially, all phonological changes that have occurred have happened via natural sound changes, but at a faster rate that than the other languages, i.e. Swahili is less similar to Proto-Bantu than the other languages.

This grammatical situation fits neatly within recent studies of the typological and sociolinguistic literature on contact: language contact results in an increased rate of change, and prolonged contact between two languages moves towards more linguistic complexity (Trudgill 2011). Prolonged contact with Arabic via the Omanis' presence in Zanzibar since the 13th century result in a strong change in the grammar of the language in comparison to other Bantu languages; however, it never blended with Arabic and became a pidgin or creole. Mufwene (2001) and Mufwene (2003) also notes the divergent behavior of Swahili when compared to other contact languages in Africa, showing that the exogamous use of Swahili has led to its adoption by the local population, which resulted in

a relatively consistent use of Swahili. From this perspective, Swahili's divergence from the other languages is attributable to the specific contact situation of prolonged bilingualism with Arabic. Crucially, none of the comparison languages have engaged in this kind of long-term bilingualism, accounting for grammatical differences between them and Swahili.

In this paper, I have compared Standard Swahili as described in Ashton (1966) to the standard varieties of several other varieties of East African Bantu languages. As just noted, standard coastal Swahili has been in long-term contact with Arabic since the 13th century, and this contact resulting in expedited change (and, at times, complexification) of several grammatical features of the standard variety. Another prediction from the literature on linguistic complexity is that simplification of grammar occurs when adult learners attempt to learn a second language. Kusters (2003a) fleshes out this claim, comparing Standard Swahili (the variety discussed in the present paper) to two other varieties of Swahili that are used as lingua francas in areas where several adult speakers of the languages speak it regularly, specifically, inland Kenyan Swahili and the Swahili spoken in the trade town of Lubumbashi in the Katanga region of the Democratic Republic of the Congo. Crucially, both of these varieties have less prestige than the coastal standard.

Kusters' findings fit the typological pattern predicted: these two lingua franca languages show several reductions in category distinctions, morphophonological complexity and a reduction of inflectional information. For reasons of space, I refer the reader to Kusters' work, but the crucial point for the current discussion is that the three varieties of Swahili are clear examples of the two kinds of second-language learning in contact areas. Standard Swahili exemplifies the effects of long-term language contact, with acquisition by young children: it has a radically divergent and at times more complex grammar than related non-contact Bantu languages. Two other varieties of Swahili that have largely been used as lingua francas by adult second-language speakers show systemic reduction in grammatical structure when compared with standard Swahili (Kusters 2003a).

Acknowledgements

I am indebted to Tony Woodbury, Pattie Epps, Peter Trudgill, and the anonymous reviews for their invaluable comments during the development of this paper. I would also like to thank the audience of the 45th Annual Conference on African Linguistics for their input and suggestions. All errors remain the fault of the author.

Abbreviations

1	first person	ASP	aspect	SG	singular
2	second person	BEN	benefactive applicative		
3	third person	PL	plural		

References

Ashton, Ethel O. 1966. *Swahili grammar*. London: Longmans.

Baker, Mark C. 2003. Linguistic differences and language design. *Trends in Cognitive Science* 7. 349–353.

Bakker, Peter. 1997. *A language of our own: The genesis of Michif, a mixed Cree-French language of the Candadian Métis*. Oxford: Oxford University Press.

Barlow, A. Ruffel. 1960. *Studies in Kikuyu grammar and idiom*. Edinburgh: William Blackwood & Sons, Ltd.

Bleek, Wilhelm. 1869. *A comparative grammar of South African languages*. London: Trübner.

Byarushengo, Ernest Rugwa, Alessandro Duranti & Larry M. Hyman. 1977. *Haya grammatical structure*. Los Angeles: University of Southern California Press.

Comrie, Bernard. 2008. Inflectional morphology and language contact, with special reference to mixed languages. In Peter Siemund & Noemi Kintana (eds.), *Language contact and contact languages*, 15–32. Amsterdam: John Benjamins.

Dahl, Östen. 2004. *The growth and maintenance of linguistic complexity*. Amsterdam: John Benjamins.

Engstrand, Olle & Abdulaziz Y Lodhi. 1985. On aspiration in Swahili: Hypotheses, field observations, and an instrumental analysis. *Phonetica* 42. 175–185.

Givón, Talmy. 1984. *Syntax: A functional-typological introduction*. Amsterdam: John Benjamins.

Givón, Talmy & Masayoshi Shibatani (eds.). 2009. *Syntactic complexity: Diachrony, acquisition, neuro-cognition, evolution*. Amsterdam: John Benjamins.

Guthrie, Malcom. 1966. *Grammaire et dictionnaire de Lingala*. Westmead, England: Gregg Press Ltd.

Guthrie, Malcom & John F Carrington. 1988. *Lingala grammar and dictionary: English–Lingala, Lingala–English*. London: Baptist Missionary Society.

Halpert, Claire. 2012. *Argument licensing and agreement in Zulu*. MIT dissertation.

Hay, Jennifer & Laurie Bauer. 2007. Phoneme inventory size and population size. *Language* 83. 388–400.

Heine, Bernd & Tania Kuteva. 2005. *Language contact and grammatical change*. Cambridge: Cambridge University Press.

Hyman, Larry M. 2003. Segmental phonology. In Derek Nurse & Gerard Philippson (eds.), *The Bantu languages*, 42–58. New York: Routledge.

Hyman, Larry M. & Francis X. Katamba. 1991. Augment in Luganda tonology. *Journal of African Languages and Linguistics* 12. 1–46.

Hyman, Larry M. & Francis X. Katamba. 1993. The augment in Luganda: Syntax or pragmatics? In Sam Mchombo (ed.), *Theoretical aspects of Bantu grammar 1*, 209–256. Stanford: CSLI.

Jerro, Kyle. 2015. Revisiting object symmetry in Bantu. In *The Selected Proceedings of the 44th Annual Conference on African Linguistics*, 130–145. Somerville, MA: Cascadilla Proceedings Project.

Katamba, Francis X. 2003. Bantu nominal morphology. In Derek Nurse & Gerard Phillippson (eds.), *The Bantu languages*, 103–120. London: Routledge.
Kimenyi, Alexandre. 1980. *A relational grammar of Kinyarwanda*. Los Angeles: University of California Press.
Kirwan, Brian Edmond Renshaw & Peter A. Gore. 1951. *Elementary Luganda*. Kampala: Uganda Bookshop.
Kusters, Wouter. 2003a. Linguistic Complexity: The Influence of Social Change on Verbal Inflection dissertation.
Kusters, Wouter. 2003b. The fate of complex languages: Classical Arabic and Old Norse in the age of globalisation. *Nordlyd* 31. 275–289.
Maddieson, Ian. 2003. The sounds of the Bantu languages. In Derek Nurse & Gerard Philippson (eds.), *The Bantu languages*, 15–41. New York: Routledge.
McWhorter, John. 2008. Why does a language undress? Strange cases in Indonesia. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds.), *Language complexity: Typology, contact, change*, 167–190. Amsterdam: John Benjamins.
McWhorter, John. 2011. *Linguistic simplicity and complexity: Why do languages undress?* Berlin: Mouton de Gruyter.
Meeussen, Achille Emile. 1967. Bantu grammatical reconstructions. *Africana Linguistica* 3. 80–122.
Miestamo, Matti. 2008. Grammatical complexity in a cross-linguistic perspective. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds.), *Language complexity: Typology, contact, change*, 23–42. Amsterdam: John Benjamins.
Miestamo, Matti, Kaius Sinnemäki & Fred Karlsson (eds.). 2008. *Language complexity: Typology, contact, change*. Amsterdam: John Benjamins.
Mufwene, Salikoko S. 2001. *The ecology of language evolution*. Cambridge: Cambridge University Press.
Mufwene, Salikoko S. 2003. Contact languages in the Bantu area. In Derek Nurse & Gerard Philippson (eds.), *The Bantu languages*, 195–208. New York: Routledge.
Myers, Scott & Megan Crowhurst. 2006. *Phonology: Case studies*. http://www.laits.utexas.edu/phonology/.
Nichols, Johanna. 2009. Linguistic complexity: A comprehensive definition and survey. In Geoffrey Sampson, David Gil & Peter Trudgill (eds.), *Language complexity as an evolving variable*, 110–125. Oxford: Oxford University Press.
Nurse, Derek. 2003. Aspect and tense in Bantu languages. In Derek Nurse & Gerard Philippson (eds.), *The Bantu languages*, 90–102. New York: Routledge.
Pinker, Steven. 1994. *The language instinct*. New York: W. Morrow.
Pinker, Steven & Paul Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13. 707–726.
Russell, Joan. 2003. *Teach yourself Swahili*. Chicago: McGraw-Hill Companies, Incorporated.
Sampson, Geoffrey. 2009. A linguistic axiom challenged. In Geoffrey Sampson, David Gil & Peter Trudgill (eds.), *Language complexity as an evolving variable*, 1–18. Oxford: Oxford University Press.

Schadeberg, Thilo C. 2003a. Derivation. In Derek Nurse & Gerard Philippson (eds.), *The Bantu languages*, 71–89. New York: Routledge.

Schadeberg, Thilo C. 2003b. Historical linguistics. In Derek Nurse & Gerard Philippson (eds.), *The Bantu languages*, 143–163. New York: Routledge.

Trudgill, Peter. 2004. Linguistic and social typology: The Austronesian migrations and phoneme inventories. *Linguistic Typology* 8. 305–320.

Trudgill, Peter. 2009. Sociolinguistic typology and complexification. In Geoffrey Sampson, David Gil & Peter Trudgill (eds.), *Language complexity as and evolving variable*, 98–109. Oxford: Oxford University Press.

Trudgill, Peter. 2011. *Sociolinguistic typology: Social determinants of linguistic complexity*. Oxford: Oxford University Press.

Zerbian, Sabine & Manfred Krifka. 2008. Quantification across Bantu languages. In Lisa Matthewson (ed.), *Quantification: A cross-linguistic perspective*, 383–415. Bingley: Emerald Publishing Group.